BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 11:16:03 ; Search time 226.903 Seconds
(without alignments)
10331.847 Million cell updates/sec

Title: US-09-807-933B-12
Perfect score: 1041
Sequence: 1 atgaagttctccatcatcgc.....ctggttgcgagcgcaagtaa 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1041	100.0	1041	21	AAA62731	Endoglucanase nucl
2	1041	100.0	1041	24	AAL43249	Phycomyces nitens
3	393.6	37.8	1017	21	AAA62729	Endoglucanase nucl
4	393.6	37.8	1017	24	AAL43247	Rhizopus arrhizus
5	382.8	36.8	1164	21	AAA62730	Endoglucanase nucl
6	382.8	36.8	1164	24	AAL43248	Rhizopus arrhizus
7	363.4	34.9	1083	21	AAA62728	Endoglucanase nucl
8	363.4	34.9	1083	24	AAL43246	Rhizopus arrhizus
9	333.8	32.1	1043	21	AAA62732	Endoglucanase nucl
10	333.8	32.1	1043	24	AAL43250	Rhizopus arrhizus
11	304.2	29.2	1017	21	AAA62726	Endoglucanase nucl
12	304.2	29.2	1017	24	AAL43244	Rhizopus arrhizus
13	289.6	27.8	1101	21	AAA62727	Endoglucanase nucl
14	289.6	27.8	1101	24	AAL43245	Rhizopus arrhizus
15	241	23.2	922	19	AAV15073	Hybrid DNA compris
16	236.6	22.7	1423	17	AAT39049	cDNA encoding cell
17	235.2	22.6	1473	12	AAQ14857	Fusarium oxysporum
18	235.2	22.6	1473	13	AAQ26407	Fusarium oxysporum
19	235.2	22.6	1473	13	AAQ26382	Endoglucanase #2.
20	235.2	22.6	1473	13	AAQ25933	Cellulase containe
21	235.2	22.6	1473	13	AAQ29935	Endoglucanase gene
22	235.2	22.6	1473	14	AAQ49942	Endoglucanase enzy
23	235.2	22.6	1473	16	AAZ60179	F. oxysporum endog
24	235.2	22.6	1473	19	AAV16103	Fusarium oxysporum
25	233.6	22.4	1473	14	AAQ41733	Dye transfer inhib
26	233.2	22.4	960	17	AAT39047	cDNA encoding cell
27	231.8	22.3	927	17	AAT39062	Chimeric endogluca
28	231.2	22.2	672	24	AAL43263	Humicola insolens
29	231.2	22.2	672	24	AAI69425	Humicola insolens
30	230.4	22.1	894	17	AAT39061	Chimeric endogluca
31	228	21.9	1174	17	AAT39050	cDNA encoding cell
32	228	21.9	1174	19	AAV39096	Monocomponent endo
33	226.2	21.7	984	19	AAV16105	Fusarium oxysporum
34	224.6	21.6	925	19	AAV15076	Hybrid DNA compris
35	222.8	21.4	928	19	AAV15074	Hybrid DNA compris
36	218.8	21.0	807	19	AAV16104	Humicola insolens
37	217.2	20.9	1058	13	AAQ26405	Humicola insolens
38	217.2	20.9	1060	12	AAQ14856	Humicola insolens
39	217.2	20.9	1060	13	AAQ26380	Endoglucanase #1.
40	217.2	20.9	1060	13	AAQ25932	Cellulase containe
41	217.2	20.9	1060	13	AAQ29934	Endoglucanase gene
42	217.2	20.9	1060	13	AAQ30067	Sequence encoding
43	217.2	20.9	1060	14	AAQ41732	Dye transfer inhib
44	217.2	20.9	1060	14	AAQ49941	Endoglucanase enzy
45	217.2	20.9	1060	16	AAZ60178	H. insolens endogl

ALIGNMENTS

RESULT 1
AAA62731
ID AAA62731 standard; DNA; 1041 BP.
XX
AC AAA62731;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 6.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Phycomyces nitens.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09826.
XX
PT Endoglucanases cf fungal origin with high activity under alkaline

```
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 128-129; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

  Query Match             100.0%;  Score 1041;  DB 21;  Length 1041;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-215;
  Matches 1041;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAGTTCTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGTTCTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCT 60

Qy         61 GAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGC 120

Qy        121 ACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCC 180

Qy        181 AACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGCCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGCCACT 240

Qy        241 ACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCT 300

Qy        301 GTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACC 360

Qy        361 ACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGC 420

Qy        421 TTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGG 480

Qy        481 GACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGT 540

Qy        541 CTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAAC 600

Qy        601 CAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGC 660

Qy        661 GCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTC 720

Qy        721 GCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAACCAC 780

Qy        781 TTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGG 840

Qy        841 AACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGC 900

Qy        901 GACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAAC 960

Qy        961 GCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAG 1020

Qy       1021 ACTGGTTGCGAGCGCAAGTAA 1041
              |||||||||||||||||||||
Db       1021 ACTGGTTGCGAGCGCAAGTAA 1041


RESULT 2
AAL43249
ID   AAL43249 standard; DNA; 1041 BP.
XX
AC   AAL43249;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Phycomyces nitens endoglucanase-related coding sequence.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Phycomyces nitens.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15057.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 81-83; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

  Query Match             100.0%;  Score 1041;  DB 24;  Length 1041;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-215;
  Matches 1041;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGAAGTTCTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGTTCTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCT 60

Qy         61 GAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGC 120

Qy        121 ACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCC 180

Qy        181 AACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGCCACT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGCCACT 240

Qy        241 ACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCT 300

Qy        301 GTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACC 360

Qy        361 ACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGC 420

Qy        421 TTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGG 480

Qy        481 GACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGT 540

Qy        541 CTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAAC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAAC 600

Qy        601 CAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGC 660

Qy        661 GCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTC 720

Qy        721 GCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAACCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAACCAC 780

Qy        781 TTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGG 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGG 840

Qy        841 AACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGC 900

Qy        901 GACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAAC 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAAC 960

Qy        961 GCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAG 1020

Qy       1021 ACTGGTTGCGAGCGCAAGTAA 1041
              |||||||||||||||||||||
Db       1021 ACTGGTTGCGAGCGCAAGTAA 1041
```

```
RESULT 3
AAA62729
ID   AAA62729 standard; DNA; 1017 BP.
XX
AC   AAA62729;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 4.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Mucor circinelloides.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09824.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 118-119; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;

  Query Match           37.8%;  Score 393.6;  DB 21;  Length 1017;
  Best Local Similarity   63.7%;  Pred. No. 8.6e-76;
  Matches  634;  Conservative    0;  Mismatches  329;  Indels   33;  Gaps    1;

Qy         46 TCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACT 105
              ||  || | ||||||   ||||||   | ||||||  || ||||||||||     |||| |
Db         55 TCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGT 114

Qy        106 GGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTAC 165
              || || ||||| ||       ||||  || || || | |   |||  |||| | |  |||
Db        115 GGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTAC 174

Qy        166 TCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACC 225
              ||||| ||| | |||      ||                                  ||
Db        175 TCTCAATGTCTTCCCGGATCCCA--------------------------------CAGT 201

Qy        226 ACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACC 285
              | ||| ||||  |   | | ||   |        ||  | ||||||  |  | || ||||||
Db        202 AACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCACC 261

Qy        286 ACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACC 345
               |||       |  ||     |||||||||| |    || |  || || |||||||| ||
```

```
Db        262 GCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAACT 321

Qy        346 ACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGC 405

Db        322 ACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACAAG 381

Qy        406 CCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAG 465

Db        382 GTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTAAA 441

Qy        466 CCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAG 525

Db        442 GCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCTCC 501

Qy        526 GATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTAC 585

Db        502 AATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTTTC 561

Qy        586 ATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCT 645

Db        562 ATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTGCT 621

Qy        646 GCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTC 705

Db        622 GCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCTTC 681

Qy        706 ACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGAT 765

Db        682 ACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCGAT 741

Qy        766 CTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGC 825

Db        742 TTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGC 801

Qy        826 TGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGC 885

Db        802 TGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGC 861

Qy        886 TCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTC 945

Db        862 TCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTC 921

Qy        946 GGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCC 1005

Db        922 AACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCT 981

Qy       1006 GAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041

Db        982 GAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
```

RESULT 4
AAL43247
ID AAL43247 standard; DNA; 1017 BP.
XX
AC AAL43247;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 4.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Mucor circinelloides.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15055.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 70-73; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;

Query Match 37.8%; Score 393.6; DB 24; Length 1017;
Best Local Similarity 63.7%; Pred. No. 8.6e-76;
Matches 634; Conservative 0; Mismatches 329; Indels 33; Gaps 1;

```
Qy         46 TCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACT 105

Db         55 TCTGCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGT 114

Qy        106 GGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTAC 165

Db        115 GGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTAC 174

Qy        166 TCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCACC 225

Db        175 TCTCAATGTCTTCCCGGATCCCA--------------------------------CAGT 201

Qy        226 ACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACC 285

Db        202 AACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCACC 261

Qy        286 ACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACC 345

Db        262 GCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAACT 321

Qy        346 ACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGC 405

Db        322 ACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACAAG 381

Qy        406 CCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAG 465

Db        382 GTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTAAA 441

Qy        466 CCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAG 525

Db        442 GCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCTCC 501

Qy        526 GATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTAC 585

Db        502 AATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTTTC 561

Qy        586 ATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCT 645

Db        562 ATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTGCT 621

Qy        646 GCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTC 705

Db        622 GCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCTTC 681
```

```
Qy        706 ACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGAT 765
              ||       ||||   |||| |||||| | || || || || ||||||||||||||  |||
Db        682 ACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCGAT 741

Qy        766 CTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGC 825
              |  || | ||||||||||||||||||| |||||||| ||||| || ||   |||||| |||
Db        742 TTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATGGC 801

Qy        826 TGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGC 885
              ||      | || |||  | |   ||  |||||||||||| ||| | ||||| ||| | |||
Db        802 TGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCAGC 861

Qy        886 TCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTC 945
              ||| | || || || | |   ||||| |      | || |||||||| || |||||||||
Db        862 TCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGATTC 921

Qy        946 GGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCC 1005
                |||||||||||| |||| |||||     | ||||||||||  ||||| || |||||
Db        922 AACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTGCT 981

Qy       1006 GAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
              ||  | | | |     | |||||||| | ||||||
Db        982 GAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1017
```

RESULT 5
AAA62730
ID AAA62730 standard; DNA; 1164 BP.
XX
AC AAA62730;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 5.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Mucor circinelloides.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09825.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 122-124; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.

XX
SQ Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;

```
  Query Match           36.8%;  Score 382.8;  DB 21;  Length 1164;
  Best Local Similarity  62.9%;  Pred. No. 1.9e-73;
  Matches  628;  Conservative    0;  Mismatches  337;  Indels   33;  Gaps    1;

Qy         44 GCTCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGA 103
              | ||| | |   |  |   ||| ||     | |||| | || || |||||||     ||||
Db        200 GTTCCTCCTCATCATCATCATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGA 259

Qy        104 CTGGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGT 163
              ||| || ||||| ||      |||| || || || | |   ||| |||| | |  |
Db        260 GTGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACT 319

Qy        164 ACTCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCA 223
              |||||||.||| | |||       ||                              ||
Db        320 ACTCTCAATGTCTTCCCGGATCCCA--------------------------------CA 346

Qy        224 CCACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCA 283
              | ||| |||  |   | | ||   |       ||  | ||||| |  | || ||||
Db        347 GTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCA 406

Qy        284 CCACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCA 343
              || |||      |  ||     |||||||||| |.  || |  || || |||||||| |
Db        407 CCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAA 466

Qy        344 CCACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACA 403
              | |||||||| | ||| ||   |||||  | ||  || |||||||     |||| ||||
Db        467 CTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACA 526

Qy        404 GCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCA 463
              |||  |||||  ||    |||||  |  |||  ||| || || || |||||||| || |
Db        527 AGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTA 586

Qy        464 AGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCA 523
              |  | ||'|||  ||||   ||||| |||||||| |||   |||||   |||||||||
Db        587 AAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCT 646

Qy        524 AGGATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCT 583
              |||||  ||      | | | || |||  || || |||||    ||| ||  | | |
Db        647 CCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTT 706

Qy        584 ACATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTG 643
              ||||||  ||  ||||||| || |||| |||||||||  || || || ||||||||||||
Db        707 TCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTG 766

Qy        644 CTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCT 703
              |||||  | | | | || ||   |  ||| ||  || || |||||||| ||  | ||||
Db        767 CTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCT 826

Qy        704 TCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATG 763
              ||||        ||||  |||| |||||| | || || || || ||||||||||||||  |
Db        827 TCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCG 886

Qy        764 ATCTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACG 823
              || |  || | |||||||||||||||||| |||||||| ||||| || ||   |||||| |
Db        887 ATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATG 946

Qy        824 GCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTA 883
              ||||      | || |||  | |   ||  |||||||||||| ||| | ||||| ||| | |
Db        947 GCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCA 1006

Qy        884 GCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGAT 943
              ||||| | || || || | |   ||||| |      | || |||||||| || |||||||
Db       1007 GCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGAT 1066

Qy        944 TCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTG 1003
              ||    |||||||||||| |||| |||||     | ||||||||||  ||||| || ||||
Db       1067 TCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTG 1126

Qy       1004 CCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
              | ||  | | | |     | |||||||  | ||||||
```

```
Db      1127 CTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1164


RESULT 6
AAL43248
ID   AAL43248 standard; DNA; 1164 BP.
XX
AC   AAL43248;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 5.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Mucor circinelloides.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15056.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 75-78; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for,
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;

  Query Match           36.8%;  Score 382.8;  DB 24;  Length 1164;
  Best Local Similarity  62.9%;  Pred. No. 1.9e-73;
  Matches  628;  Conservative    0;  Mismatches  337;  Indels   33;  Gaps    1;

Qy          44 GCTCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGA 103
               | ||| | |    |  |   ||| ||    | |||| | || || |||||||     ||||
Db         200 GTTCCTCCTCATCATCATCATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGA 259

Qy         104 CTGGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGT 163
               ||| || ||||| ||        ||||  || || || | |   |||  |||| | |  |
Db         260 GTGGACCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACT 319

Qy         164 ACTCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCCAAGACCACCACCACCACCA 223
               |||||||(||| | |||     ||                                  ||
Db         320 ACTCTCAATGTCTTCCCGGATCCCA------------------------------CA 346

Qy         224 CCACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCA 283
               | ||| ||||  |   | | ||    |        ||  | ||||||  |  | || ||||
Db         347 GTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTACCAAGACATCTACTACCA 406

Qy         284 CCACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCA 343
               || |||      |  ||     |||||||||| |    || |  || || |||||||| |
```

```
Db         407 CCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGACAA 466

Qy         344 CCACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACA 403
               | |||||||| | ||| ||    |||||  | ||  || |||||||     |||| ||||
Db         467 CTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTACA 526

Qy         404 GCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCA 463
                  ||| ||||| ||     |||||  |  |||  ||| || || || ||||||||  || |
Db         527 AGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTGTA 586

Qy         464 AGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCA 523
               |  | ||||||  ||||   ||||| |||||||| |||   |||||  ||||||||||
Db         587 AAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGCCT 646

Qy         524 AGGATGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCT 583
                 |||||  ||        | | | || |||  || || |||||   ||| ||  | |  |
Db         647 CCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGGTT 706

Qy         584 ACATGTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTG 643
                |||||| ||  |||||||  || |||| |||||||| || || || ||||||||||||
Db         707 TCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGCTG 766

Qy         644 CTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCT 703
               |||||  | | | | || ||    |  ||| ||  || || |||||||| ||  | ||||
Db         767 CTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGACCT 826

Qy         704 TCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATG 763
               ||||        ||||  |||| |||||| | || || || || ||||||||||||  |
Db         827 TCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGGCG 886

Qy         764 ATCTCAGCACCAACCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACG 823
               || |  || | |||||||||||||||||| |||||||| ||||| || ||    ||||||
Db         887 ATTTAGGCTCTAACCACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGGTATCTTCAATG 946

Qy         824 GCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTA 883
               ||||     | || ||| | |    || |||||||||| ||| | ||||| ||| | |
Db         947 GCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATATGGTGGTGTCA 1006

Qy         884 GCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGAT 943
               ||||| | || || || | |   |||||  |     | || |||||||| || |||||||
Db        1007 GCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTGTAAATGGAGAT 1066

Qy         944 TCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTG 1003
               ||   |||||||||||| |||| |||||     | ||||||||||  ||||| || ||||
Db        1067 TCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGTTACCTGTCCTG 1126

Qy        1004 CCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
               | ||  | | | |     | ||||||||  | ||||||
Db        1127 CTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1164


RESULT 7
AAA62728
ID   AAA62728 standard; DNA; 1083 BP.
XX
AC   AAA62728;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 3.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Rhizopus oryzae.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
```

```
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09823.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs  -
XX
PS   Claim 44; Page 113-115; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

  Query Match             34.9%;  Score 363.4;  DB 21;  Length 1083;
  Best Local Similarity   61.3%;  Pred. No. 2.9e-69;
  Matches  652;  Conservative    0;  Mismatches  381;  Indels   30;  Gaps    3;

Qy          9 CTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAG 68
              |||| |      |  || |||  | ||| | |       |  | ||||||||||| ||
Db         21 CTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATGGCCCATGCTGCTGAATGTAG 80

Qy         69 CCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGG 128
              | |  ||   ||| |||||||| ||||  |||  ||| || ||||||||     || ||
Db         81 CAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGAATCTGG 140

Qy        129 CTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCA 188
              ||  || || || | |       ||||    |  ||||| || ||| | ||||| ||  |
Db        141 CTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAATGAAAA 200

Qy        189 AGT-----------CCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGC 236
              |           | |   |  | ||| |||||||| ||    |  |||||    |
Db        201 CCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAAGAAGAC 260

Qy        237 CACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGC 296
              || || || || | | |     |  |||||||   |      || | |  | ||||   |
Db        261 TACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 320

Qy        297 CCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCAC 356
              |||   |  |||  ||| |||| ||||||    || | ||||  |||||||  || ||
Db        321 TACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCACTACTAC 380

Qy        357 CA---------CCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCC 407
              |         | |||   || || |  || | || ||| |    ||   ||||  | |
Db        381 TAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTACTCCGC 440

Qy        408 CATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCC 467
              | ||||||||   ||| || || |||   || || ||||||||||||| || |||||
Db        441 TGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCC 500

Qy        468 CTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGA 527
              ||||||   |||  ||| |||||  ||| ||    |||||   | |||||  ||||||
Db        501 TTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGA 560

Qy        528 TGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACAT 587
              ||||   |  | |   | |  | |    || | ||| || || || ||       |||||
Db        561 TGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAGCTACAC 620

Qy        588 GTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGC 647
              || |||||||| || || |||||||| |  |||||||||||||||||||||| |||||
Db        621 CTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGC 680

Qy        648 CAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCAC 707
              |   | | |   | |   || | |  || || | ||||| ||| || || |||||
Db        681 TTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCAC 740

Qy        708 CAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCT 767
              |   ||||| |||   || |||||| | || || || || |||||||| |||  ||| ||
Db        741 CTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCT 800

Qy        768 CAGCACCAAC---------CACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTT 818
              || | |||         |||||||| ||||| |||||||| |||||||| ||   ||
Db        801 TGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTA 860

Qy        819 CAACGGCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGG 878
              ||| || ||      | || |||    |   ||||||||| |||||||  | || |||||
Db        861 CAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGG 920

Qy        879 TATTAGCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTG 938
              | ||   |||   || || ||    || |||||| |      | || ||||||||  |||||
Db        921 TGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTG 980

Qy        939 GAGATTCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTG 998
              |||||||| |||||||| |||||||| ||||||    | |||| |||    |||||  ||
Db        981 GAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTG 1040

Qy        999 CCCTGCCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
              ||    |  |||| ||||||| |||| ||      | || |||
Db       1041 TCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083
```

```
RESULT 8
AAL43246
ID   AAL43246 standard; DNA; 1083 BP.
XX
AC   AAL43246;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 3.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15054.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 65-68; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
```

CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

Query Match 34.9%; Score 363.4; DB 24; Length 1083;
Best Local Similarity 61.3%; Pred. No. 2.9e-69;
Matches 652; Conservative 0; Mismatches 381; Indels 30; Gaps 3;

```
Qy    9 CTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAG 68
Db   21 CTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATGGCCCATGCTGCTGAATGTAG 80

Qy   69 CCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGG 128
Db   81 CAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGAATCTGG 140

Qy  129 CTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCA 188
Db  141 CTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAATGAAAA 200

Qy  189 AGT----------CCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGC 236
Db  201 CCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAAGAAGAC 260

Qy  237 CACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGC 296
Db  261 TACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 320

Qy  297 CCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCAC 356
Db  321 TACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCACTACTAC 380

Qy  357 CA---------CCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGGCTACAGCCC 407
Db  381 TAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTACTCCGC 440

Qy  408 CATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCC 467
Db  441 TGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCC 500

Qy  468 CTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGA 527
Db  501 TTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGA 560

Qy  528 TGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACAT 587
Db  561 TGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAGCTACAC 620

Qy  588 GTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGC 647
Db  621 CTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGC 680

Qy  648 CAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCAC 707
Db  681 TTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCAC 740

Qy  708 CAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCT 767
Db  741 CTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCT 800

Qy  768 CAGCACCAAC---------CACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTT 818
Db  801 TGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTA 860

Qy  819 CAACGGCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGG 878
Db  861 CAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGG 920

Qy  879 TATTAGCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTG 938
Db  921 TGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTG 980

Qy  939 GAGATTCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTG 998
Db  981 GAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTG 1040

Qy  999 CCCTGCCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
Db 1041 TCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083
```

RESULT 9
AAA62732
ID AAA62732 standard; DNA; 1043 BP.
XX
AC AAA62732;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 7.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 132-134; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;

Query Match 32.1%; Score 333.8; DB 21; Length 1043;
Best Local Similarity 66.5%; Pred. No. 7.2e-63;
Matches 530; Conservative 0; Mismatches 252; Indels 15; Gaps 3;

```
Qy  260 CCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCTGTCACCAC---CACCAAGG 316
Db  236 CCCACAAGAAGACCACGACCGCTGCCCACAAGAAGACCACGACCGCCGCTCACAAGAAGA 295

Qy  317 CCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCA 376
Db  296 CTACGACCGCTCCCGCCAAGAAGACCACGACCGTCGCCAAGGCTTCGACTCCGTCCAACT 355

Qy  377 CCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTC 436
```

```
Db       356 CGAGCAGCTCGTCTTCGGGAAAGTACAGCGCTGTCAGCGGTGGCGCTAGCGGCAACGGCG 415

Qy       437 GCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTG 496

Db       416 TCACTACCCGCTACTGGGACTGCTGCAAGGCTTCGTGCTCGTGGCCCGGCAAGGCTAACG 475

Qy       497 TAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGTCT---CGGTTCCGATG 553

Db       476 TCAGCTCGCCTGTCAAGTCCTGCAACAAGGACGGCGTCACCGCTCTTAGCGACTCCAACG 535

Qy       554 TCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTG 613

Db       536 CCCAGTCCGGCTGCAACGGCGGCAACTCCTACATGTGCAACGACAACCAGCCATGGGCTG 595

Qy       614 TCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTG 673

Db       596 TCAACGACAACCTTGCTTACGGTTTCGCTGCCGCTGCCATTAGCGGCGGTGGCGAGAGCC 655

Qy       674 CATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGT 733

Db       656 GCTGGTGCTGCTCCTGCTTCGAGCTCACCTTCACCTCCACCAGCGTTGCTGGCAAGAAGA 715

Qy       734 TTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAA---------CCACTTTG 784

Db       716 TGGTCGTCCAGGTCACCAACACTGGCGGTGACCTTGGCAGCTCGACCGGTGCCCACTTCG 775

Qy       785 ATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACA 844

Db       776 ATCTCCAGATGCCCGGCGGCGGCGTCGGCATCTTCAACGGATGCTCGTCCCAGTGGGGCG 835

Qy       845 CCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACA 904

Db       836 CTCCCAACGACGGCTGGGGCTCGCGCTACGGCGGCATCAGCTCCGCCAGCGACTGCTCGT 895

Qy       905 AGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTG 964

Db       896 CCCTCCCCAGCGCCCTCCAGGCCGGCTGCAAGTGGCGCTTCAACTGGTTCAAGAACGCCG 955

Qy       965 ACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTG 1024

Db       956 ACAACCCGTCCATGACCTACAAGGAGGTCACCTGCCCCAAGGAGATCACCGCTAAGACCG 1015

Qy      1025 GTTGCGAGCGCAAGTAA 1041

Db      1016 GATGCTCGCGCAAGTAA 1032
```

```
RESULT 10
AAL43250
ID   AAL43250 standard; DNA; 1043 BP.
XX
AC   AAL43250;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related codon-optimised DNA sequence.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
OS   Synthetic.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15052.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Example 10; Page 84-86; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;
```

```
  Query Match           32.1%;  Score 333.8;  DB 24;  Length 1043;
  Best Local Similarity   66.5%;  Pred. No. 7.2e-63;
  Matches  530;  Conservative    0;  Mismatches  252;  Indels   15;  Gaps    3;

Qy       260 CCACCACCAAGGCCACCACCACCACCACCACCAAGGCCCCTGTCACCAC---CACCAAGG 316

Db       236 CCCACAAGAAGACCACGACCGCTGCCCACAAGAAGACCACGACCGCCGCTCACAAGAAGA 295

Qy       317 CCACTACTACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCA 376

Db       296 CTACGACCGCTCCCGCCAAGAAGACCACGACCGTCGCCAAGGCTTCGACTCCGTCCAACT 355

Qy       377 CCACCTCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTC 436

Db       356 CGAGCAGCTCGTCTTCGGGAAAGTACAGCGCTGTCAGCGGTGGCGCTAGCGGCAACGGCG 415

Qy       437 GCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTG 496

Db       416 TCACTACCCGCTACTGGGACTGCTGCAAGGCTTCGTGCTCGTGGCCCGGCAAGGCTAACG 475

Qy       497 TAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAGCCGTCT---CGGTTCCGATG 553

Db       476 TCAGCTCGCCTGTCAAGTCCTGCAACAAGGACGGCGTCACCGCTCTTAGCGACTCCAACG 535

Qy       554 TCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTG 613

Db       536 CCCAGTCCGGCTGCAACGGCGGCAACTCCTACATGTGCAACGACAACCAGCCATGGGCTG 595

Qy       614 TCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTG 673

Db       596 TCAACGACAACCTTGCTTACGGTTTCGCTGCCGCTGCCATTAGCGGCGGTGGCGAGAGCC 655

Qy       674 CATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGT 733

Db       656 GCTGGTGCTGCTCCTGCTTCGAGCTCACCTTCACCTCCACCAGCGTTGCTGGCAAGAAGA 715

Qy       734 TTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAA---------CCACTTTG 784

Db       716 TGGTCGTCCAGGTCACCAACACTGGCGGTGACCTTGGCAGCTCGACCGGTGCCCACTTCG 775

Qy       785 ATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACA 844

Db       776 ATCTCCAGATGCCCGGCGGCGGCGTCGGCATCTTCAACGGATGCTCGTCCCAGTGGGGCG 835

Qy       845 CCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACA 904

Db       836 CTCCCAACGACGGCTGGGGCTCGCGCTACGGCGGCATCAGCTCCGCCAGCGACTGCTCGT 895

Qy       905 AGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTG 964
```

```
Db        896 CCCTCCCCAGCGCCCTCCAGGCCGGCTGCAAGTGGCGCTTCAACTGGTTCAAGAACGCCG 955

Qy        965 ACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTG 1024

Db        956 ACAACCCGTCCATGACCTACAAGGAGGTCACCTGCCCCAAGGAGATCACCGCTAAGACCG 1015

Qy       1025 GTTGCGAGCGCAAGTAA 1041

Db       1016 GATGCTCGCGCAAGTAA 1032


RESULT 11
AAA62726
ID   AAA62726 standard; DNA; 1017 BP.
XX
AC   AAA62726;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 1.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Rhizopus oryzae.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09821.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 104-105; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

  Query Match           29.2%;  Score 304.2;  DB 21;  Length 1017;
  Best Local Similarity   59.9%;  Pred. No. 1.8e-56;
  Matches  616;  Conservative    0;  Mismatches  353;  Indels   60;  Gaps    4;

Qy          25 GCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAG 84

Db          37 GCTCTCGCCCTCGGTACTGAAATGGCCTCTGCTGCTGAATGTAGCAAATTGTATGGTCAA 96

Qy          85 TGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGT 144

Db          97 TGTGGTGGTAAGAACTGGAATGGCCCTACTTGTTGTGAATCTGGATCCACCTGTAAA--- 153

Qy         145 GCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCC 204

Db         154 ---GTAAGCAACGATTACTACTCTCAATGTCTTCCCTCTGGAAGCAGTGGCAATAAATCT 210

Qy         205 AAGACCACCACCACCACCACCACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACC 264

Db         211 TCTGAAAGTGCTCACAAGAAGACTACCACTGCTGCTCACAAGAAGACTACTACCGCTGCT 270

Qy         265 ACCAAGGCCACCACCACCACCACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACT 324

Db         271 CATAAA----------------------------------------AAGACTACCACT 288

Qy         325 ACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCC 384

Db         289 GCTCCTGCTAAGAAGACTACAACTGTTGCCAAAGCTTCCACCCCTTCTAACTCTAGCTCT 348

Qy         385 TCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACC 444

Db         349 AGCTCCAGCGGCAAATATTCCGCTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACTACT 408

Qy         445 CGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAG 504

Db         409 CGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGTTCT 468

Qy         505 CCTGTACTCACCTGTGCCAAGGATGGTGTCA---GCCGTCTCGGTTCCGATGTCCAGAGC 561

Db         469 CCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGCCCAAAGT 528

Qy         562 GGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTGTCAATGAC 621

Db         529 GGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAACCTTGGGCTGTAAACGAC 588

Qy         622 GACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGC 681

Db         589 AACCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGGTGC 648

Qy         682 TGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTC 741

Db         649 TGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTTGTC 708

Qy         742 CAGGTCACCAACACCGGTGATGATCTCAGCACC---------AACCACTTTGATTTGCAG 792

Db         709 CAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTGCAA 768

Qy         793 ATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACACCAACACC 852

Db         769 ATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGTTCCAGCCAATGGGGTGCTCCCAAT 828

Qy         853 GATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACAAGCTTCCT 912

Db         829 GACGGTTGGGGCTCAAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTTCCT 888

Qy         913 ACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTGACAACCCA 972

Db         889 TCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAACCCA 948

Qy         973 GAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTGGTTGCGAG 1032

Db         949 AGCATGACTTACAAGGAAGTTACCTGTCCTAAGGAAATCACCGCCAAGACAGGTTGTTCA 1008

Qy        1033 CGCAAGTAA 1041

Db        1009 AGAAAATAA 1017


RESULT 12
AAL43244
ID   AAL43244 standard; DNA; 1017 BP.
XX
AC   AAL43244;
XX
DT   22-AUG-2002  (first entry)
```

XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 1.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15052.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Example 10; Page 56-58; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

Query Match 29.2%; Score 304.2; DB 24; Length 1017;
Best Local Similarity 59.9%; Pred. No. 1.8e-56;
Matches 616; Conservative 0; Mismatches 353; Indels 60; Gaps 4;

```
Qy         25 GCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAGCCAAGGCTATGGCCAG 84
Db         37 GCTCTCGCCCTCGGTACTGAAATGGCCTCTGCTGCTGAATGTAGCAAATTGTATGGTCAA 96

Qy         85 TGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGGCTTCACCTGTGTAGGT 144
Db         97 TGTGGTGGTAAGAACTGGAATGGCCCTACTTGTTGTGAATCTGGATCCACCTGTAAA--- 153

Qy        145 GCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCAAGTCCAGGGTAACCCC 204
Db        154 ---GTAAGCAACGATTACTACTCTCAATGTCTTCCCTCTGGAAGCAGTGGCAATAAATCT 210

Qy        205 AAGACCACCACCACCACCACCACCAAGGCTGCCACTACCACCAAGGCTCCTGTCACCACC 264
Db        211 TCTGAAAGTGCTCACAAGAAGACTACCACTGCTGCTCACAAGAAGACTACTACCGCTGCT 270

Qy        265 ACCAAGGCCACCACCACCACCACCACCAAGGCCCCTGTCACCACCACCAAGGCCACTACT 324
Db        271 CATAAA----------AAGACTACCACT 288

Qy        325 ACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCACCACCTCC 384
Db        289 GCTCCTGCTAAGAAGACTACAACTGTTGCCAAAGCTTCCACCCCTTCTAACTCTAGCTCT 348

Qy        385 TCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACC 444
Db        349 AGCTCCAGCGGCAAATATTCCGCTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACTACT 408

Qy        445 CGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAG 504
Db        409 CGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGTTCT 468

Qy        505 CCTGTACTCACCTGTGCCAAGGATGGTGTCA---GCCGTCTCGGTTCCGATGTCCAGAGC 561
Db        469 CCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGCCCAAAGT 528

Qy        562 GGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTGTCAATGAC 621
Db        529 GGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAACCTTGGGCTGTAAACGAC 588

Qy        622 GACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGC 681
Db        589 AACCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGGTGC 648

Qy        682 TGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTC 741
Db        649 TGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTTGTC 708

Qy        742 CAGGTCACCAACACCGGTGATGATCTCAGCACC---------AACCACTTTGATTTGCAG 792
Db        709 CAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTGCAA 768

Qy        793 ATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACACCAACACC 852
Db        769 ATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGTTCCAGCCAATGGGGTGCTCCCAAT 828

Qy        853 GATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACAAGCTTCCT 912
Db        829 GACGGTTGGGGCTCAAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTTCCT 888

Qy        913 ACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTGACAACCCA 972
Db        889 TCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAACCCA 948

Qy        973 GAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTGGTTGCGAG 1032
Db        949 AGCATGACTTACAAGGAAGTTACCTGTCCTAAGGAAATCACCGCCAAGACAGGTTGTTCA 1008

Qy       1033 CGCAAGTAA 1041
Db       1009 AGAAAATAA 1017
```

RESULT 13
AAA62727
ID AAA62727 standard; DNA; 1101 BP.
XX
AC AAA62727;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 2.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX

```
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09822.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 108-110; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences.(see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;

  Query Match           27.8%;  Score 289.6;  DB 21;  Length 1101;
  Best Local Similarity   64.3%;  Pred. No. 2.6e-53;
  Matches  471;  Conservative    0;  Mismatches  249;  Indels   12;  Gaps    2;

Qy        322 ACTACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCACCACC 381
              || ||||| |||||     || | |   | |  || || ||        ||  | | | |
Db        370 ACGACTACTACCACTGCTCCCGCTAAGGAAATTACAACTACTGCCAAAGCTTCAAACTCT 429

Qy        382 TCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACT 441
              ||    || | |       |||  |    | ||||||||   |||||| ||||||  ||||
Db        430 TCTAACTCTAGCGGCAAATACTCCATTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACT 489

Qy        442 ACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACT 501
              || || || |||||||||| ||| |||| ||   |||| ||| |||||   ||| | |
Db        490 ACTCGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGT 549

Qy        502 AAGCCTGTACTCACCTGTGCCAAGGATGGTGTCA---GCCGTCTCGGTTCCGATGTCCAG 558
                |||||     |||||  ||| ||||||||||     || |   |     | |||||||
Db        550 TCTCCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGTCCAA 609

Qy        559 AGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTGTCAAT 618
              || || ||   ||| ||  |    |||||||| || ||||||||||| |||| ||¶ ||
Db        610 AGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAGCCTTGGGCTGTAAAC 669

Qy        619 GACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTC 678
              ||  | |||||||| ||||||||||||||     || || | |  |||| |||   |
Db        670 GATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGG 729

Qy        679 TGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTC 738
              |||||    |||| ||| ||||| |||||    ||  |||| ||||| |||||| | ||
Db        730 TGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTT 789

Qy        739 GTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAAC---------CACTTTGATTTG 789
              |||| ||||| ||||| |||| ||||||  || ||              |||||||| |||
Db        790 ATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTG 849

Qy        790 CAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACACCAAC 849
              || |||||||| |||||||| ||    ||||| || |||       || |||    |  |
Db        850 CAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAGCAATGGGGTGCTCCC 909

Qy        850 ACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACAAGCTT 909
              |  || || |||||  |  | || || ||||||   |||   || || |||   |  |||
Db        910 AATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTT 969

Qy        910 CCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTGACAAC 969
              ||| ||    | || |||||||| || |||||||||   ||||||||||||||||| |||
Db        970 CCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAAC 1029

Qy        970 CCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTGGTTGC 1029
              |||    | || | |||||  ||||| || ||    || |||| ||||||||| |||||
Db       1030 CCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACCGCCAAGACAGGTTGT 1089

Qy       1030 GAGCGCAAGTAA 1041
                | || |||
Db       1090 TCAAGAAAATAA 1101


RESULT 14
AAL43245
ID   AAL43245 standard; DNA; 1101 BP.
XX
AC   AAL43245;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 2.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15053.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp -
XX
PS   Disclosure; Page 60-63; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;

  Query Match           27.8%;  Score 289.6;  DB 24;  Length 1101;
  Best Local Similarity   64.3%;  Pred. No. 2.6e-53;
  Matches  471;  Conservative    0;  Mismatches  249;  Indels   12;  Gaps    2;

Qy        322 ACTACTACCACCACCAAGACCACCACCAAGACCACCACCACCAAGGCTGCCACCACCACC 381
              || ||||| |||||     || | |   | |  || || ||        ||  | | | |
Db        370 ACGACTACTACCACTGCTCCCGCTAAGGAAATTACAACTACTGCCAAAGCTTCAAACTCT 429

Qy        382 TCCTCTTCCAACACTGGCTACAGCCCCATTTCTGGTGGCTTCTCTGGAAACGGTCGCACT 441
              ||    || | |       |||  |    | ||||||||   |||||| ||||||  ||||
Db        430 TCTAACTCTAGCGGCAAATACTCCATTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACT 489

Qy        442 ACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACT 501
              || || || |||||||||| ||| |||| ||   |||| ||| |||||   ||| | |
Db        490 ACTCGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGT 549

Qy        502 AAGCCTGTACTCACCTGTGCCAAGGATGGTGTCA---GCCGTCTCGGTTCCGATGTCCAG 558
                |||||     |||||  ||| ||||||||||     || |   |     | |||||||
Db        550 TCTCCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGTCCAA 609
```

```
Qy        559 AGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGACAACCAGCCCTGGGTTGTCAAT 618
              || || ||    ||| ||  |    |||||||| || ||||||||||| |||| ||| ||
Db        610 AGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAGCCTTGGGCTGTAAAC 669

Qy        619 GACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGGTAGCGCCGGTGCCTCTGCATTC 678
              ||  | ||||||||| ||||||||||||||     || || | |  ||||  |||    |
Db        670 GATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGG 729

Qy        679 TGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGCTGTCGCTGGCAAGAAGTTTGTC 738
              |||||    |||| ||| ||||| |||||     ||  |||| ||||| |||||| | ||
Db        730 TGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTT 789

Qy        739 GTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAAC---------CACTTTGATTTG 789
               |||| ||||| ||||| |||| ||||||   || ||               |||||||| |||
Db        790 ATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTG 849

Qy        790 CAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCAGTGGAACACCAAC 849
              || |||||||| |||||||| ||    ||||| || |||     || |||    |   |
Db        850 CAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAGCAATGGGGTGCTCCC 909

Qy        850 ACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGAGTGCGACAAGCTT 909
              |. || || ||||| |  | || || |||||   |||   || || |||   |  |||
Db        910 AATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTT 969

Qy        910 CCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAAGAACGCTGACAAC 969
              ||| ||    | || |||||||| || |||||||||   ||||||||||||||||| |||
Db        970 CCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAAC 1029

Qy        970 CCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGCCAAGACTGGTTGC 1029
              |||    | || | |||||  ||||| || ||    || ||||  |||||||| |||||
Db       1030 CCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACCGCCAAGACAGGTTGT 1089

Qy       1030 GAGCGCAAGTAA 1041
                | || |||
Db       1090 TCAAGAAAATAA 1101
```

RESULT 15
AAV15073
ID AAV15073 standard; DNA; 922 BP.
XX
AC AAV15073;
XX
DT 19-MAY-1998 (first entry)
XX
DE Hybrid DNA comprising a family 45 cellulase core region.
XX
KW Cellulase core region; isolation; microorganism; identification;
KW hybrid DNA; ss.
XX
OS Chimeric - Cylindrocarpon sp.
OS Chimeric - Humicola insolens.
XX
FH Key   Location/Qualifiers
FT CDS   2..922
FT   /*tag= a
FT   /product= "hybrid gene construct"
XX
PN WO9743409-A2.
XX
PD 20-NOV-1997.
XX
PF 12-MAY-1997; 97WO-DK00216.
XX
PR 10-MAY-1996; 96DK-0000562.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Dalboge H, Diderichsen B, Kauppinen S, Sandal T;
XX
DR WPI; 1998-008878/01.
DR P-PSDB; AAW44270.
XX
PT Isolating novel DNA sequences from microorganisms - without the need
PT for culturing the microorganism
XX
PS Example 3; Page 53-54; 72pp; English.
XX
CC The present sequence represents a novel hybrid gene construct
CC from an example of the present invention. The present invention
CC describes a novel method for providing a novel DNA sequence encoding a
CC polypeptide from a microorganism with an activity of interest. The
CC method comprises: (i) PCR amplification of the DNA with PCR primers with
CC homology to (a) known gene(s) encoding a polypeptide with an activity of
CC interest; (ii) linking the obtained PCR product of a 5' structural gene
CC sequence and a 3' structural gene sequence; (iii) expressing the
CC resulting hybrid DNA sequence; (iv) screening for hybrid DNA sequences
CC encoding a polypeptide with the activity of interest or a related
CC activity; and (v) isolating the hybrid DNA sequence identified in step
CC (iv). This method provides for identification and isolation of sequences
CC from microorganisms without having to cultivate and isolate the
CC microorganism.
XX
SQ Sequence 922 BP; 166 A; 325 C; 257 G; 174 T; 0 other;

  Query Match   23.2%; Score 241; DB 19; Length 922;
  Best Local Similarity 63.1%; Pred. No. 7.9e-43;
  Matches 392; Conservative   0; Mismatches 220; Indels   9; Gaps   1;

```
Qy        417 TGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCCCTCTTGCGC 476
              ||  |||  |||    | ||  |  | ||  | ||||||||||| || ||||||||||| |
Db         61 TGCTTTCGCTGCTGATGGCAGGTCCACGAGGTACTGGGATTGTTGTAAGCCCTCTTGCTC 120

Qy        477 CTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGATGGTGTCAG 536
              |||||  ||  ||||| || || |      || || || ||||| | |||| | |    |
Db        121 CTGGGGCGACAAGGCCTCGGTCAGCGCCCCCGTCCTGACCTGCGACAAGAACGACAACCC 180

Qy        537 CCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACATGTGCAATGA 596
              |     ||   || | | |  |||||||||| |||||||    ||||||  ||||   |
Db        181 CATCTCCGACGCCAACGCCGTGAGCGGTTGCAACGGCGGCACTTCCTACACCTGCAGCAA 240

Qy        597 CAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGCCAGTCTCGG 656
              ||||   || |||| |||||| ||| |||| ||||| || || || ||| |||  |||
Db        241 CAACTCCCCGTGGGCTGTCAACGACAACCTCGCCTATGGCTTTGCCGCTACCAAGCTCTC 300

Qy        657 TAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCACCAACACTGC 716
              | | | |   |  ||    | ||||| |  || ||||  || ||||| || |      |
Db        301 TGGAGGCTCCGAGTCCAGCTGGTGCTGTGCTTGCTACGCTCTCACCTTTACGACTGGCCC 360

Qy        717 TGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCTCAGCACCAA 776
              ||     |||||||   | ||||| ||| ||||||||||||| |  |||||| ||   ||
Db        361 CGTGAAGGGCAAGACCATGGTCGTACAGTCCACCAACACCGGAGGCGATCTCGGCGAGAA 420

Qy        777 CCACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTTCAACGGCTGCCAGTCCCA 836
              |||||| ||| | |||||||||||||| |||||||||   |||  |||||||   |||||
Db        421 CCACTTCGATCTCCAGATGCCCGGCGGCGGTGTCGGCATCTTTGACGGCTGCAGCTCCCA 480

Qy        837 GTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGGTATTAGCTCTATTTCAGA 896
              ||||               |    | | |||||||  || ||||| ||     ||       ||
Db        481 GTGG---------GGTGGCCTCGGCGGTGCTCAGTACGGCGGCATCTCGTCGCGAAGCGA 531

Qy        897 GTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTGGAGATTCGGATGGTTCAA 956
              |||||||   | ||    | | |  |||  || ||| | ||| | | | ||  ||||||||
Db        532 CTGCGACAGCTTCCCCGAGCTGCTCAAGGACGGCTGCTACTGGCGCTACGACTGGTTCAA 591

Qy        957 GAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTGCCCTGCCGAGATCATTGC 1016
              |||||| ||||| ||     ||| ||||       ||    ||||| |||||| || | ||
Db        592 GAACGCCGACAATCCGAGCTTCAGCTTCCGTCAGGTCCAGTGCCCAGCCGAGCTCGTCGC 651

Qy       1017 CAAGACTGGTTGCGAGCGCAA 1037
                || || |||   |||||
Db        652 TCGCACCGGATGCCGCCGCAA 672
```

Search completed: June 17, 2003, 11:49:21

Job time : 232.07 secs